

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2002, 06:43:24 ; Search time 38.27 seconds
(without alignments)
1076.834 Million cell updates/sec

Title: US-09-652-292-2
Perfect score: 2765
Sequence: 1 MCHSPVPLCASVLLGL.....CHRONSTGTPYSRIEISAAS 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:.*
1: pirl:.*
2: pirl:.*
3: pirl:.*
4: pirl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626.5	22.7	457	2 E70070	metabolite transpo
2	595	21.5	461	2 D70073	metabolite transpo
3	594.5	21.5	580	2 D86426	hypothetical prote
4	567	20.5	521	2 C84864	probable membrane
5	545	19.7	580	2 D84772	probable sugar tra
6	541	19.6	582	2 F71431	hypothetical prote
7	521.5	18.9	433	2 G86812	D-xylose proton-sy
8	520.5	18.8	557	2 T38125	myo-inositol trans
9	508	18.4	493	2 A85433	sugar transporter
10	499	18.0	454	2 F65079	galactose-proton s
11	497.5	18.0	547	2 A48442	membrane transport
12	497	18.0	511	2 A84537	probable sugar tra
13	497	18.0	606	2 T27072	hypothetical prote
14	496	17.9	464	2 F85951	galactose-proton s
15	495	17.9	472	2 B26430	L-arabinose isomer
16	495	17.9	472	2 B85936	hypothetical prote
17	495	17.9	560	2 T51495	sugar transporter-
18	491	17.8	523	2 S25015	monosaccharide tra
19	489.5	17.7	472	2 S47089	arabinose-proton s
20	488.5	17.7	464	2 F69587	L-arabinose transp
21	488.5	17.7	522	2 E86246	glucose transport
22	484.5	17.5	511	2 H84536	probable sugar tra
23	484.5	17.5	522	2 S12042	glucose transport
24	483	17.5	516	2 T12199	monosaccharid tran
25	482	17.4	508	2 T05156	probable glucose t
26	480.5	17.4	549	2 T14606	probable sugar tra
27	473.5	17.1	575	2 T43400	myo-inositol trans
28	471.5	17.1	639	2 T23658	hypothetical prote
29	467	16.9	523	2 T10122	hexose transport p

30	466.5	16.9	508	2 G84564	probable sugar tra
31	464.5	16.8	468	2 S10014	glucose transport
32	460	16.6	487	2 E96782	hypothetical prote
33	460	16.6	491	2 A26430	xylose transport p
34	460	16.6	491	2 B86096	xylose-proton symp
35	455.5	16.5	502	2 B70845	probable sygar tra
36	452.5	16.4	482	2 B69803	metabolite transpo
37	451.5	16.3	613	2 T27077	hypothetical prote
38	450	16.3	473	2 G89789	sugar transporter
39	450	16.3	547	2 C84593	probable sugar tra
40	438	15.8	584	2 S69555	myo-inositol trans
41	435	15.7	490	2 T14545	probable sugar tra
42	434.5	15.7	513	2 T01506	probable hexose tr
43	430.5	15.6	522	2 T10150	sugar transport pr
44	426.5	15.4	504	2 T00450	probable monosacch
45	426	15.4	472	2 T35662	probable sugar tra

ALIGNMENTS

RESULT 1

E70070

metabolite transport protein homolog ywtg - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: E70070

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabis, C.; Ferrari,
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portote
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033

A:Accession: E70070

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-457 <KUN>

A:Cross-references: GB:299122; GB:AL009126; NID:q2636029; PIDN:CAB15600.1; PID:g26361

A:Experimental source: strain 168

C:Genetics:

A:Gene: ywtg

C:Superfamily: glucose transport protein

Query Match 22.7%; Score 626.5; DB 2; Length 457;

Best Local Similarity 30.1%; Pred. No. 2.1e-34;

Matches 160; Conservative 85; Mismatches 179; Indels 107; Gaps 8;

Oy 17 LGLNFTGYELAVISGALLPLQDFGLSCLEQDFLVGSLGALLGALLVGLFDYGRKQ 76

Db 15 LGLALYGTGTGVSIGAILPMKELGLNATEGLVSSVLGAILGSGAAGKLTDRGRKK 74

Oy 77 ATILGNLVLAGSLTGLAGSLAWLVGRVAVFAISLSSMACCIYVSELVGRQGVLV 136

Db 75 ATMAAALFICIGLGVAPLNTGVNLFRIILGLAVGTSTTVPLVLSLAPKHKRGALS 134

Oy 137 SYEAGITVGIILSTALYNALACTPWGRHMFQWATAPAVLQSLSLPLP-----AGT 189

Db 135 SUNQLMITVGIILSVIVNYIFADAE--AWRMILGLAAVPSLLILIGILFMPESPRLWLTNG 193

Oy 190 DETATHTKDLIPLOG-----GEAPKLGPRPRYSFLDLFRARDNMGRHTTVGLGL 238

Db 194 EESKAKKILKLEKLGTKDIDQEIHDKEAKQDEG----GLKELFDPW--VRPALIAGLGL 247

QY 239 VLFOQLTQCPNVLCYASTIFSSVFGHGSSAVLASVGLGAVKVAATLTAMGLVDRAGRA 298
 Db 248 AFLQOFGINTIIYAPKFTNVGP-GNSASILGTGTGVTGNVLMVAIKIIDIKGRKP 306
 QY 299 LILACALMALSVSGIGLVSAFVPMDSGSCSLAVPNATGOTGLPGDGLLODSSLPPR 358
 Db 307 LULFGNAGVTSILVALVNLFF----- 329
 QY 359 TNEDQREPLSTAKTKPHRSGDPSAPPRALSSALPCPPLPARGHALLRWATLLCLMV 418
 Db 330 -----NWTPAA-----SWTVCILGV 345
 QY 419 FVSASFSGFPGVTNLVLSIYVEIRGRAFAFCNSFNMAANLFISSLFDLIGTIGLSWT 478
 Db 346 FIVFVAVSGVPMVWMLPELPLHVRGIGTGVSTLMLHVGVLIVSLYPIYPMERIGISYL 405
 QY 479 FLITGLTAVLGLGFIYLPVETKQSLAEIDQOQKRRFTLSFGHRSNTG 529
 Db 406 FLIVAAIGIMAFVRFKVTETKRSLEETEQDLDRKNGOGGAAGKQOTVG 456

RESULT 2

D70073
 metabolite transport protein homolog yxcC - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence-revision 05-Dec-1997 #text-change 20-Jun-2000
 C:Accession: D70073
 R:Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertelli, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chedid, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinol, Y., M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamamoto, K.; Yata, K.; Yoshida, K.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D70073
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <KUN>
 A:Cross-references: GB:299124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g2636527
 A:Experimental source: strain 168
 C:Genetics:
 C:Superfamily: glucose transport protein

Query Match 21.5%; Score 595; DB 2; Length 461;
 Best Local Similarity 29.0%; Pred. No. 2.7e-32;
 Matches 151; Conservative 95; Mismatches 162; Indels 112; Gaps 11;
 QY 17 LGLTFGYELAVISCALLPLQDLGSLCEQBEFLVGLSLLGALLASLVGGFLIDYGRKQ 76
 Db 16 LGLLYGDTGVIGISALLFINNDPLTLTTECLVYVSMLLGAIFGSALSGTCSRWGRK 75
 QY 77 AILGSNLVLLAGSLTIGLAGSLAWLVLCRAVVGFAISLSSMACCIYVELSGPRQGVLV 136
 Db 76 VVFLSIIFIIGLACAFSQIGMLIASRVILGLAVGGSTALPVVYLSEMAPTKIRGTG 135
 QY 137 SLYEAGITVGLLYALNVALAGTPW-GWRHMFGWATAPAVLOSLSLLFLP-----A 187
 Db 136 TMNMLIVTIGLLAYIVNYLF--TPFEARWVGLAAVPAVLLIGIAGFPSPRWLVKR 193
 QY 188 GTDETA-----TH-----KDLIFLOGEAPKLGFCRPRYSFLDLPRDRNRGRTTVG 235
 Db 188 GTDETA-----TH-----KDLIFLOGEAPKLGFCRPRYSFLDLPRDRNRGRTTVG 235

Db 194 GSEEEARRIMNITHDPKDIEMELAEKQGEAEK-----KETTGLGVKAK-WIRPMLLIG 246
 QY 236 LGLVLFQOLTQCPNVLCYASTIFSSVFGHGSSAVLASVGLGAVKVAATLTAMGLVDRAG 295
 Db 247 VGLAIFQOAVGINTVIYAPKFTNVGP-GTSASALGTMGIGILNVIMCITAMILIDRVG 305
 QY 296 RRALLAGCALMALSVSGIGLVSAFVPMDSGSCSLAVPNATGOTGLPGDGLLODSSLPP 355
 Db 306 RKLLIWSGVITLSLAAL-----SGVLLTGLS- 334
 QY 356 IPRNEDQREPLSTAKTKPHRSGDPSAPPRALSSALPCPPLPARGHALLRWATLLC 415
 Db 335 -----ASTA-----WMTVVF 344
 QY 416 LMVFSAFSGFPGVTNLVLSIYVEIRGRAFAFCNSFNMAANLFISSLFDLIGTIGL 475
 Db 345 LGVIVFYQATWGPVWMLPELFPSPKARGAATGFTTLVLSAANLIVSLVFLPMLSAMGI 404
 QY 476 SWTFLLYGLTAVLGLGFIYLPVETKQSLAEIDQOQKRR 515
 Db 405 AMVFMVSVICLLSFFFAFYMPETKGSLEIEEASLKKR 444

RESULT 3
 D86426
 hypothetical protein AAG30955.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text-change 31-Mar-2001
 C:Accession: D86426
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alor, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurco, J.S.; Maiti, R.; Marzla Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: D86426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <STO>
 A:Cross-references: GB:AE005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 21.5%; Score 594.5; DB 2; Length 580;
 Best Local Similarity 29.6%; Pred. No. 3.7e-32;
 Matches 173; Conservative 94; Mismatches 217; Indels 101; Gaps 16;
 QY 4 SPVPLPCASVSLGGLTFGYELAVISCALLPLQDLF---GLSCLEQEFVLGSLIGALL 60
 Db 25 NPVYLRFAFSAG-IGGLFGYDTGTVISGALLYIRDFKSDVDRNTWLOEMIVSNVAGAI 83
 QY 61 ASLVGGFLIDYGRKQAILGSNLVLLAGSLTIGLAGSLAWLVLCRAVVGFAISLSSMAC 120
 Db 84 GAAIGGWANDKLGRSAILMADFLLGATIMAAAPNSLLVGRVFGVGLGVGNASMTAP 143
 QY 121 IYVSELVGPQRQGVLSLYEAGITVGLLYALNVALAGTPWGRHMFGWATAPAVLOS 180
 Db 144 LYISEASPARKALYSTNGFLITGTGQFSLYNLAFTDVTGTWRMLGTAGIPALLQV 203
 QY 181 SLFLP-----ACTDETA-----THKDLIP---LOGEAPKLG 211
 Db 204 LMFTLPSRWLYRKGRSEEAAILRRIYSAEDVEQEIHALKDSVETEIEGSSSEKI--- 261
 QY 212 GRPRYSFLDLPRDRNRGRTTVVGLVLFQOLTQCPNVLCYASTIFSSVFGHGSSAVL 271
 Db 262 -----NWIKLCAKTVRRG-LTAGVGLVQQFVGINTVWYISPTIVQLAGFASNRALL 315

QY 272 ASVGLGAVKVAATLTAMGLVDGRRAILLACALMALSVSGIGLVFAVPMDSGSPCLA 331
 Db 316 LSLVATAGNAGSIISIFIDIRGRKKLLI-----ISLFCV-IISUGI----- 357
 QY 332 VFNATG-----QTGLPGDSGL-----LDSSLPPIPTNEDQREPISTAKTKP----- 376
 Db 358 ---LTVGYEAATHAPALISLETORFNINISCPDYKSAMNTNWDCTCLKASPSGCGYCS 414
 QY 377 -----HPRSAPP-----RIALSSALPGPPLPARGHALLRWTALLCLMVFVS 421
 Db 415 SPIGKEHPCACWISDSVDKLDCHNENRLMYTRGCP-----SNFGWALLGLGLYII 465
 QY 422 AFSPGCGPVTWLVSEIYVPEIRGRAFAFCNSFNNAANLISLSFLDLIGTIGLSTWTL 481
 Db 466 FSPGNGVTPWLVNSEIYVPEIRGCGIAATANISNLIVAQSFSLTEAIGTSWTELI 525
 QY 482 YGLTAVLGLGFIYLVPEPKGOSLAIEDQOFKRRFTLSFGHQN 526
 Db 526 FGVISVIALLFVWVCVPETKGMPEIEKMLERRSMERKFKKS 570
 RESULT 4
 G84864
 probable membrane transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G84864
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: G84864
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-521 <STO>
 A:Cross-references: GB:AE002093; NID:g2289003; PIDN:AB864332.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g43330
 A:Map position: 2
 C:Superfamily: glucose transport protein

Query Match 20.5%; Score 567; DB 2; Length 521;
 Best Local Similarity 27.8%; Pred. No. 2.2e-30;
 Matches 150; Conservative 85; Mismatches 183; Indels 122; Gaps 6;
 QY 17 LGGLTFGYELAVISGALLPLQLDL-----GLSCLEQFVLGSLIGALLA 61
 Db 40 IGLLFGDTGTVISGALLYIKDDFEVVKOSFLQVNVSSFTSSKLETIVSMALVGAMIG 99
 QY 62 SLVGGFLIDCYGRKQAILGSLNVLGSLTGLAGSLAWLVGRVAVGPAISLSSMACCI 121
 Db 100 AAGGWINDYIGRKKATLADYVFAAGAVMAAAPDPVYLISGRLLVGLGCVASVAPV 159
 QY 122 YVSELVGRQGVLSVLEAGITVGLISLYALNYALAGTPGWRHMFHGWATAPVQLQSL 181
 Db 160 YIAEASPEVRGGLVSTNVLMTGQFLSYLNSAFTVPGTWRWMLGVSVPVAVIOFIL 219
 QY 182 LLFLP-----ACTDETATHKDLIPLOGGEAPKLGPRPRYSFLDLF 222
 Db 220 MLFNPESRWLMKRRKAEAIQVLTADYISREDEIDHLSAAEEEEKKRTKTVGLDVF 279
 QY 223 RARDNMRGRTTVGLGLVLFQQLTGOPNVLCYASTIFSSVGFHGGSSAVLASVGLGAVKA 282
 Db 280 RSKE-LRLAFLAGAGLQAFQQTGINTVYISPTIVOMAGFHSNQLALFLSLIVAAMNA 338
 QY 283 ATLTAMGLVDGRRAILLAG-----CALMALSVSGIGLVFAVPMDSGSPCLAVPNATG 337
 Db 339 GTVVGIVFYDHCGRKKLALLSFLGVIISLLISVSFF-----K 376
 QY 338 QTGLPGDSGLQDSSLPIPTNEDQREPISTAKTKPXRSGDPSAPPRLLSALPG 397

Db 377 QSETSSDGG----- 385
 QY 398 PPLPARGHALLRWTALLCLMVFVSFAFGPVPVTLVLSLSEIYVPEIRGRAFAFCNSFNNA 457
 Db 386 -----LYGWLAVLGLALYIVFFAPGMPVPMVNSEIYPOQYRGICGGNSATVNI 436
 QY 458 ANLFTLSFLDLIGTIGLSTWTFLLYGLTAVLGLGFIYLVPEPKGOSLAIEDQOFKRRF 517
 Db 437 SNLVAQTLETTAEAAAGTGTFTLILAGIAVLAVIFVIVFVPEQGLTFSEVEQIKERAY 496
 RESULT 5
 D84772
 probable sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: D84772
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: D84772
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <STO>
 A:Cross-references: GB:AE002093; NID:g4263781; PIDN:AAD15441.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g35740
 A:Map position: 2

Query Match 19.7%; Score 545; DB 2; Length 580;
 Best Local Similarity 28.6%; Pred. No. 7e-29;
 Matches 164; Conservative 99; Mismatches 205; Indels 106; Gaps 18;
 QY 4 SPVPLPLCASVSLGLLTFGYELAVISGALLPLQLDLFG---LSCLEQFVLGSLIGALL 60
 Db 23 TPVIMRLALSAG-IGGLFGYNTGTAGALLYIKKEEFGVDNKTWLIQIIVSMTVAGAIV 81
 QY 61 ASLVGGFLIDCYGRKQAILGSLNVLGSLTGLAGSLAWLVGRVAVGPAISLSSMACC 120
 Db 82 GAAGGWINDYIGRKKATLADYVFAAGAVMAAAPDPVYLISGRLLVGLGCVASVAPV 141
 QY 121 YVSELVGRQGVLSVLEAGITVGLISLYALNYALAGTPGWRHMFHGWATAPVQLQSL 180
 Db 142 LYISEMSPARIGALVSTNGLITGQFLSYLNLAFVHTPGTWRWMLGVSVAIPAIIQFC 201
 QY 181 SLFLPAG-----TDETATHKDLI-----PLOGGEA-----PKLGPRP 214
 Db 202 LMLTLPESRWLYRNDKKAESRDILRIYPAENVEAIAALKESVRAETADEIIG---- 257
 QY 215 RYSEFLDFR-ARDN--MRGRTTVGLGLVLFQQLTGOPNVLCYASTIFSSVGFHGGSSAV- 270
 Db 258 -HTFSKLGALSNPVVRHGLAGITVQVQAQPVGINTVYISPTIQLQFAGYASNTAMA 316
 QY 271 --LASVGLGAVKVAATLTAMGLVDGRRAILLAGSLAWLVGRVAVGPAISLSSMACC 328
 Db 317 LALITSLANV---GSVSMFMVDYGRKKLMI-----ISMFGII-----T 354
 QY 329 CLAVPNATGQTGLPGDGLLQDSSLPIPTNEDQRE-----PILS---TAKTKPH-- 377
 Db 355 CLVILAAVFNPA-----SHAPKIDK--RDSRNFANATCPAFAPPTASRSPSNW 403
 QY 378 -----PRSCDPSAPPRLLSALPGPPLPARGHALLR-----WTALLCL 416
 Db 404 NCMKCIQYDCGFCGNGAQETAPGACTVQSDMKALCHSKGRTEFFKDCGCKSKGYLAIVFL 463
 QY 417 MVFVSFAFGPVPVTLVLSLSEIYVPEIRGRAFAFCNSFNNAANLISLSFLDLIGTIGL 476
 Db 464 GLYIIVAFMGCTVPMVNSEIYPLRYGLAGGIAAVSNMNSLVSEITFLTITNAVGS 523

Badcock, K

adcock, K.; Churcher, C.M.; Wood, V.; Barrall, P.C.; Paterson, M.

RESULT 8
T36125
myo-inositol transporter 2 - fission yeast (Schizosaccharomyces pombe)
C:\Species\Schizosaccharomyces pombe
C:\Data:03-Dec-1999#sequences pombe
C:\Data:03-Dec-1999 #text_change 21-Jan-2000
Accession: T36125
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1997

A:Accession: Z21772

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-557 <BAD>

A:Cross-references: EMBL:Z95334; PIDN:CAB08597.1; GSPDB:GN00066; SPDB:SPAC20G8.03

A:Experimental source: strain 972h-; cosmid c20G8

C:Genetics:

A:Gene: SPDB:SPAC20G8.03

A:Map position: 1

C:Superfamily: maltose transport protein MAL61

Query Match 18.8%; Score 520.5; DB 2; Length 557;
Best Local Similarity 27.1%; Pred. No. 2.8e-27;
Matches 144; Conservative 90; Mismatches 189; Indels 109; Gaps 10;

```
QY 12 ASVSLGLTFCYELAVISGALLPLQLDPG--LSCLEQEFVLSLLGALLASLVGGFLI 69
DB 85 SAVAGISGLLFGTGVISGALVLSGDLGHVLSCKRELITSATSAALISATSGWLA 144
QY 70 DCYGRKQATLGSNLVLLAGSLTGLAGSLAWLVGRVGVFAISLSSMACCIYVSELVGP 129
DB 145 DWGKRKLLCADAFIVIGSVTMAASRNVAHVGRFIVGYGIGLTSLLVPMYTELAPA 204
QY 130 RORGVLVSYEAGITVIGLLSYALNYALAGTGWGRHMFGNATAPAVLQSLSLFLPAGT 189
DB 205 RLURGLVITYVYFRTGGQLIAYSLNAAFEHVHGWRIMFGIGAAPALGQLISLFWTPESP 264
QY 190 DETATH---KDLPILOGGEAPKLGPCRPRY-----SFLDLF 222
DB 265 RYLLRHNVHVKYKILSRTHPEAKPAETIAIKVSLIOGVKVDFFPEGNKFOHFFHSLKVL 324
QY 223 RARDNMRGRTTVGLGLVLFQOLTGPONVLCYASTIFSSVFGHGSSAVLASVGLGAKVA 282
DB 325 TVPSNRRS-LFICGFLOWFOQFSGTNAIOYFSAIFQSIVGFGF---RNSISVIVGATNFV 380
QY 283 AFLTAMGLVDGRALLAGCALMALSVSGIGLVSAFVPMDSGSPCLAVPNATGOTGLP 342
DB 381 FTIVAFMFIDIRGRIRILCTSAVM---IAGLALCAIAYHF-----LP 420
QY 343 GDSGLLDSSLLPIPTNEDREPILSTAKTKPHRSGDPSAPPLALSSALPGPPLA 402
DB 421 AD-----TQNTNSG----- 430
QY 403 RGHALLRMTALLCLMVFSAFSGFGPVTWLVSEIYVEIRGRAPAFCSNFWAANLFI 462
DB 431 -----WQYVVLASIIIFLASYASGIGNIPW-QQAELEFVNEVRALGAGFTAINWVGNLII 484
QY 463 SUSFLDLICTIGLSWTFLLYGLTAVLGLGFIYLFVETKQSLAEIDQOFQX 514
DB 485 SASFLTMESITPTGTFTALFAGFCFVGLVTSYFTYPELAGMSIENIHKLEK 536
```

RESULT 9

A:Accession: A85433

A:Title: sugar transporter like protein [imported] - Arabidopsis thaliana

A:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C:Accession: A85433

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A:Reference sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488

A:Accession: A85433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:NC_001269; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g336670

A:Map position: 4

C:Superfamily: glucose transport protein

Query Match 18.4%; Score 508; DB 2; Length 493;

Best Local Similarity 25.5%; Pred. No. 1.7e-26;

Matches 140; Conservative 83; Mismatches 194; Indels 132; Gaps 9;

```
QY 11 CASVSLGGTLTFCYELAVISGALLPLQLDPGLSCLSQEFLVGLSLLGALLASLVGGFLID 70
DB 20 CAIVASIVSIIFGYDTGVMSGANVFIEEDKTNDVQTEVLGTILNLCALVGSLLAGRTSD 79
QY 71 CYGRKQATLGSNLVLLAGSLTGLAGSLAWLVGRVGVFAISLSSMACCIYVSELVGP 130
DB 80 IIGRRITIVLASILFPMGLSILMGWGNYPVLLSGRGTAGLVGFALMVAIVPYSAELATAS 139
QY 131 ORGVLSVLYEAGITVIGLLSYALNYALAGTPW--GWRHMFGNATAPAVLQSLSLFLPAG 188
DB 140 HRGLASLPHLCISIGLLGYIVNYPFESKPLMHIGWRLMLGIAAVPSLVLAFLGILKMPES 199
QY 189 -----TDETHAKDLPILOGGEAPKLGPCRPRYSFLDLFRAD-----NMRGRT 232
DB 200 PRLIMQGRLEKEKEILELVSN-----SPEAELEFQDIKAAAGIDPKCYDDVVKMEGKK 254
QY 233 TVG-----LGLVLFQOLTGPONVLCYASTIFSSVFGHGSSAV 270
DB 255 THGEGVWKEILRPDPVRRVLLTALCIIHFQHASGIEAVLLYGPRIKXKAGITTKDKLF 314
QY 271 LASVGLGAVKVAATLTMGLVDRAGRALL--AGCALMALSVSGIGLVSAFVPMDSGSPS 328
DB 315 LVTVIGVIMKTTFTATLLEDKVRKRLLTSTVSGNVIALTMLGPGLT----- 363
QY 329 CLAVPNATGOTGLPGDSSLLQDSSLLPIPTNEDREPILSTAKTKPHRSGDPSAPPR 388
DB 364 --MAQNAGGK----- 371
QY 389 LALSSALPGPPLPARGHALLRMTALLCLMV--FVSASFSGFGPVTWLVSEIYVPEIRG 445
DB 372 -----LAWALVLSIVAAYSFAFFSIGLIGPITWVYSEVFPKLRA 412
QY 446 RAPAFCSNFWAANLFIISLISFLDLICTIGLSWTFLLYGLTAVLGLGFIYLFVETKQSL 505
DB 413 OGASLGVAVNRVNVATVSNHSLTSLSAITGGGAFFMAGVAAVANWVFFFLPETKQKSL 472
QY 506 AEIDQOFQX 514
DB 473 EEIEALFOR 481
```

RESULT 10

F65079

galactose-proton symport (galactose transporter) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 24-Sep-1999

C:Accession: F65079

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: F65079

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-464 <BLAT>

A:Cross-references: GB:AE0000377; GB:U00096; NID:g2367178; PIDN:AACT5980.1; PID:g17893

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: galP

C:Superfamily: glucose transport protein

Query Match 18.0%; Score 499; DB 2; Length 464;

Best Local Similarity 27.2%; Pred. No. 6.2e-26;

Matches 141; Conservative 95; Mismatches 183; Indels 100; Gaps 10;

```

QY 14 VSLGGTGFYELAVISGALLPLQLODFGLSCLEQEFVLGSLLLCALLASLVGGFLDCVG 73
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 21 LAALAGLLFGULDIGVJAGALPFIADPEQIITSHTQEWVSSMMFGAAGVAGSGLSKLG 80
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 74 RKQALCSNLVLLAGSTLGLAGSLAWLVJGRVAFGAFISLSSMACCIYVSELVGRQRG 133
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 81 RKSLMIGAILFVAGSLFSAAPNVEVILSRVLLGLAVGVASYTAPLYLSEIAPKIRG 140
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 134 VLVSLEYAGITVGTLLSYALNYALAGTPhGWRHMFHGWATAPAVLQSLFLFPAQTDETA 193
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 141 SMISYQIMTITGILGALSDFSTF-GAWRRMLGVIIIPAILLLIGVFELPDSPRMPFA 199
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 194 THKDLIPLOQ-----GGEAPK-LCPGRP-----RYSFLDLFRANDNRGRTTVCLGLV 239
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 200 AKREFYDAERVLRLRDTSAEARELDETRIESLOVQSGWALFKENSNFRRAVFLGVLLQ 259
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 240 LFOOLTGOPNVLYCASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDYRAGRRL 299
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 260 VMOOFTGMNVIMYAPKIFELAGYTWNTTQOMGTIVGLTNVLATFIAGLVDRGRKPT 319
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 300 LLACCALMALSVSGIGLVSPVMDSPGSLAVPNATGOTGLPGDGLLQDSSLPIPRPT 359
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 320 LTJGLFLVNA---AGMGVLG-----335
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 360 NEDQREPILSTAKTKTPHRSRGGPSAPPRIALSSALPGPPLPARGHALLRWTTALLCLMV 419
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 336 -----TWHHGTIISPSA-----OYFAIAMLLMF 358
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 420 VSAPFSGFGBVTVLWLSVEIYVBEIRGAPAF-CN-SFNMAANLFISLSFLDGLIGTGLSW 477
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 359 IVGFAMASAGPLIWLVCSEIQ--LKGRDFGTCSTATNVIANNIVCAFTLWNLTLGNAN 416
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
QY 478 TFLLYGLTAVILGFGFIYFLVPETKQSGSLAEIQQQOKRR 516
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :
Db 417 TFWYIYAALNVFLTLTLVLPVETPKHVSLEHIRENLMKGR 455
   : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | :

```

```

RESULT 11
A48442
membrane transport protein (clone pL5H) - Leishmania donovani
C:Species: Leishmania donovani
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: A48442
R:Langford, C.K.; Ewbank, S.A.; Hanson, S.S.; Ullman, B.; Landfear, S.M.
Mol. Biochem. Parasitol. 55, 51-64, 1992
A:Title: Molecular characterization of two genes encoding members of the g1
A:Reference number: A48442; MUID:93063058
A:Accession: A48442
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-547 <L>A>
A:Cross-references: GB:M85072; NID:g159301; PID:AAA29230.1; PID:g159302
A:Note: sequence extracted from NCBI backbone (NCBIN:118159, NCBIPI:118162)
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein

```

Query Match .18.0%; Score 497.5; DB 2; Length 547;
Best Local Similarity 26.4%; Pred. No. 9.4e-36;
Matches 146; Conservative 83; Mismatches 190; Indels 135; Gaps 15;

QY	10	LCASVLLGGLTFGYELAVISGALLPQLDFGLS---CLBQEFVLGSLLAGALLASLVGGF	67
Db	7	LCAA---LGGFLFGYDTGVINAALFQMKDHFGESEHSWQYALIVATAIACAFVGAFISGF	63
QY	68	LIDCTGRKQAILGSLNVLVLGASITLGLAGSLAWLVLGRAVVGFAISLSLMACCIYVSEIV	127
Db	64	ISAAFGRRPCTAVADALFVIGSVLGAAPNVEVLVSRVIVGLAIGSSATPIVLAETV	123
QY	128	GPQRQGVLSLYEAGITTVCGILLS---YALNYALAGTPCWQRHFWGATAPAVLQSLSLFLF	184
Db	124	SPKHGATIVLNNLFTGCGQFVAAGTAIMVVFTEKNTGMRVAIGTALPAVVQAFCLLF	183

QY	185	-----LPAGTDETA-----HKDLPILO--GGEAPKLGPCGRYSFLDLFRADNM	228
Db	184	FLPESPWLLSKGHADRAKAVADFEVDLCEFOGDELPSV-----RIDYRPLM-ARD-M	236
QY	229	RGRITVGLGVLFOOLTQOPNWLYCYASTIFSSVGHGSSAVLASVGLGAVKVAATLTAM	288
Db	237	RFRVVISGLGLIQOQFSGINTINTYSSVILYDAGFRAIMPVVLISPLAFMNAFLTAVAI	296
QY	289	GLVDRAGRRAALLA--G-CALMALSVSGIGL-----VSFAVPMDSGSLAVPNATGQTG	340
Db	297	FTVDPRRRRMLLSVFGCLVLLVIAIIFFGITRISYSV-----	337
QY	341	LPQSGLLQDSSLPPIPRTNEDQREPTLSTAKTKPHPRSGDPSAPPRLASLPGPLP	400
Db	338	--GGGLF-----	342
QY	401	PARGHALLRWTALLCLMVFSAPSGFGPPTWLVLSIYVPEITRGRAFPCNSFNWAHL	460
Db	343	-----LALLAVFLAYAPGICGPWYIMGEIFPTRLTSAASVATMANNGAVN	390
QY	461	FISLSFLDLIGTIGLSMTFLLYGLTAVLGGFYVLKVPETKGOSLAEIDQOQKR-----	515
Db	391	LVSQVFPIIMGAIGVGGTFTIISGLMALGCFIVFFFAVEYKGTULTSQIDMNFKRAGLPP	450
QY	516	RFTLSFGHRQNSTG	529
Db	451	RF-----HEGESG	459

RESULT 12

RESULT
A84537

probable sugar transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

```
C:\Program Files\Microsoft Office\Office11\MSOFTPLT.DLL
```

```
C:\Date: 02-Feb-2001 #sequence_revlsion 02-Feb-2001 #text_change 16-Feb-2001
```

C;Accession: A84537

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fuhl, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; Tallon.

euss, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A; Reference number: A84420; MUID: 20083487

A; Accession: A84537

A;Status: preliminary

A; Molecule type: DNA

A;Residues: 1-511 <STO>

A; residues: I-JII <SIU>
A; Cross-references: GB:

A; Cross-References: GA;
C; Genetics:

C;GENETICS:
A:Gene: AT2g16130

A:Gene: AT2G16130
A:Map position: 2

A;Map position: 2
C:Superfamily: glucose

C;superfamily: glucose

Query Match

Query Match	18.0%	Score 497;	DB 2;	Length 511;
Best Local Similarity	26.5%;	Pred. No. 9.4e-26;		
Matches 149; Conservative	80;	Mismatches 197;	Indels 136;	Gaps 11;

Qy	11	CASVSLUGGTFGYELAVISGALLPQLDFGLSCLEQEFVLGSLLIGALLASLVGGFLD	70
		: : : : : : : : : :	
Dd	29	CALLASMTSIIILYDIGVMGSAAIKIKDKLSDVOLEIIMGLINLYLSIGSGAAGRTSD	88
		: : : : : : : : : :	

QY 71 CYGRQQAILGSNVLVLGAGSLTGLAGSLAWLVLGRAVVGFAISLSMSNACCIIYVSELVGP 130
 89 WIGRRYTVLAGFFFCGALMGENTNYPIMVREFGVAGICVGYAMMIADPVTYTEVDAS 148

QY 131 ORGLVLSLYEAGITVGILLSYALNVALACTP - WQWRHMFPGWATPAVLQSLSLFLPAG 188
 149 SRGFTSSSEPTETNGLICLVSVSVFETKPSHCTPBMVCTCAUDCT : : : : :
 db 149 SRGFTSSSEPTETNGLICLVSVSVFETKPSHCTPBMVCTCAUDCT : : : : : 200

QY	189	-----TDETATHKD-----	LIPOGGEAPK	208
		: : :	:	
DB	200	-----	-----	-----

203 FAREYRQGRGDRFAYLDRKTSNINEEAKSKUNDAKNAVGLPDDMTDDVIWVP-----NK 262

Qy 209 LQCPRRYSFLDLPARDNMRRTTVGLGLVLPFOOLTQOPNVLCYASTIFSSVGFHGGSS 268
 Db 263 KSAGGVWMDL-LVPTPSVRIHILACIHFSSQASGIDAVVLYSPITFISRAGLSKND 321
 Qy 269 AVLASVGLGAVKVAATLTAMGLVDRAGRRALLAGCALMALSVSGIGLVSVFAVPMDSGPS 328
 Db 322 QLLATVAVGVVKTFLFVVGTCVLDVDFRGRALLLTSMGMFFSLTALG-TSLTV-IDRNP- 378
 Qy 329 CLAVPNATGOTGLPGDGLQDSSLPPIPTNEDOREPILSTAKTKPHPRSGDPSAPPR 388
 Db 379 -----GQT----- 381
 Qy 389 LALSALGPPLPARGHALLRWT---ALLCLMVPVAFSPFGPVTWLVLSIYVEIRG 445
 Db 392 -----LKAWGLAVTMTVFVATFSLGAGPVTWYASIEFVRRA 422
 Qy 446 RAFAPCSFNMAANLFLSFLDLTIGTIGLSWTFLLYGLTAVLGLGYFLFVPETKGSGL 505
 Db 423 QGASGLVNLRLMSGIIGTIGTIGLSKGLTIGGAFLLFAGVAAVWVFFTFPLPETRGVPL 482
 Qy 506 AEIDQOF-----QKRRFTLSFG 522
 Db 483 BEIESLFGYSANKKNVMSKG 504

RESULT 13
 T27072
 hypothetical protein Y51A2D.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T27072
 R:McMurray, A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: T20307
 A:Accession: T27072
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-606 <WIL>
 A:Cross-references: EMBL:AL021497; PIDN:CAA16400.1; GSPDB:GN00023; CESP:Y51A2D.4
 A:Experimental source: clone Y51A2D
 C:Genetics:
 A:Gene: CESP:Y51A2D.4
 A:Map position: 5
 A:Introns: 4/3; 76/1; 144/1; 226/3; 400/2; 446/1; 562/3
 C:Superfamily: glucose transport protein

Query Match 18.0%; Score 497; DB 2: Length 606;
 Best Local Similarity 26.5%; Pred. No. 1.1e-25;
 Matches 160; Conservative 94; Mismatches 247; Indels 102; Gaps 15;

Qy 3 HSPV--LPLCASVSLGLTFFGYELAVISGALLPLQLDFGLSCLE---QEFVLGSLLL 56
 Db 17 HTPKGLFVYILAAASVIGGFLGYDTSVVSANLYMPDAPGLKPMDFWQEVLSISPG 76
 Qy 57 GALLASLVGGFLIDCYGRKQKQAILGSNLVLLAGSLTGLAGSLAMVLGRVAVGFAISLSS 116
 Db 77 MAVGSLSGTSDDYIGRRKVLGASAIPTIGALVCAASVKNIMLLVGRVLLGIAIGFAS 136
 Qy 117 MACCIYSELVGPORGVLSVLYEAGITVGLLSY----ALNYALAGTPW--GWRHMGW 170
 Db 137 MIVPVYGETAPTHVRGMLVAFAALMISFGOVVANIITGAFSYI---DPYVNGVRLMFAF 193
 Qy 171 ATAPAVLQSLILFLP-----AGTDEATHTKDLIPLQGGEPAPKLGPRPRYSFLDLF- 222
 Db 194 AAVFSIQVCFMFLPETPRMLYENGFTETREVLEKYNGDKWV-----EYEMAEIIA 248
 Qy 223 -----RARDNMRRTT-----VGLGLVLPFOOLTQOPNVLCYASTIFSS 260
 Db 249 FNEQAKENEAHAGSPVIWRLTKTPHYLKACFIGSMQLQAQOOLAGINTILYTTADIIRS 308
 Qy 261 VGFHGGSAVLASVGLGAVKVAATLTAMGLVDRAGRALLAGCALMALSVSGIGLVSPA 320

Db 309 SCISNNHTTINISVLLSLCNFICGPFVPMNSLIEKVGRIIFESCGLVLSLVFIGVAFIL 368
 Qy 321 VPMDSGPS-----CLAVPN-----ATGOTGLPGDGLQDSSLPPI 356
 Db 369 VNHDSAAATLPAHQYGSNFSNYPDAKGCMAYSNCDYCVTTDACGFCHDANTKQGYCLP-- 426
 Qy 357 PRTNEDOREPILSTAKTKPHPRSGDPSAPPRLLALSSALPGPLPARGHALLRWTALLCL 416
 Db 427 --AGFDNDEVYSTGCTNSNGSIANNFKWYKCYCDTKYLLPIACG----- 472
 Qy 417 MVFVSFAFSGFGPVTWLVLSIYVEIRGAFACFNSFNMAANLFLSFLDLTIGTIGLS 476
 Db 473 -VYLLTFSSGFTSLPVLNLSFEYPMARSTCVASTSNVFNLIILATVLSLTQVIGY 531
 Qy 477 WTFLYGLTAVLGLGYFLFVPETKGSGLAEIDQOF-----QKRRFTLS-FGHQ 525
 Db 532 GAFWLYAGLTVIAFIPLVFPETKGYSTEEVEMLFMKNKKORREASRRRETVEYRSRM 591
 Qy 526 NST 528
 Db 592 NST 594

RESULT 14
 F85951
 galactose-proton symport of transport system [imported] - Escherichia coli (strain O1
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
 C:Accession: F85951
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: F85951
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-464 <STO>
 A:Cross-references: GB:A8005174; NID:gl2517486; PIDN:AAG58074.1; GSPDB:GN00145; UWGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: galP
 C:Superfamily: glucose transport protein

Query Match 17.9%; Score 496; DB 2: Length 464;
 Best Local Similarity 27.2%; Pred. No. 9.9e-26;
 Matches 141; Conservative 95; Mismatches 183; Indels 100; Gaps 10;

Qy 14 VSLGLGLTFGYELAVISGALLPLQLDFGLSCLEQEFVLGSLILGALLASLVGGFLIDCYG 73
 Db 21 LAALAGLLFGIDIGVTAGALPTIADEFQITSHTEWVSSMMFMGAAGVAVGSGWLSFKLG 80
 Qy 74 RKQAILGSNLVLLAGSLTGLAGSLAMVLGRVAVGFAISLSSMACCIYVSELVGPQRG 133
 Db 81 RKSLMIGAILFVAGSLFSAAPNPEVLTSRVLLGLAVGVSATYAPLVLSEIAPEKIRG 140
 Qy 134 VLVSVEAGITVGLLSVLYALNVALAGTPMGRHMEGWATAPAVLOSLSLFLPAGTDETA 193
 Db 141 SMISWQLMITTIGLAVLSUDTAFSYT-GAWRMGLGVIIIPAILLLIGVFFLPDPSRWF 199
 Qy 194 THKDLIPLQ-----GGEAPK-LGPGRP----RYSFLDLFRARDNMRRTTVGLGLV 239
 Db 200 AKRRFVDAERVLRLRDTSAEAKRELDREISLVQKSGWALFKENSNFRRAVFLGLVQ 259
 Qy 240 LFOOLTQOPNVLCYASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAGRALL 299
 Db 260 VMOQFTGMNVIYAPKIPFELAGYNTTBOHMGTVITGLTNVLTATFATGLVDRMGRKPT 319
 Qy 300 LLAGCALMALSVSGIGLVSVFAVPMDSGLAVPNATGOTGLPGDGLQDSSLPPIPT 359
 Db 320 LTLGLFLVMA---AGMGLG----- 335

QY 360 NEDQREPISTAKTKPHRPSGDPSPAPRLALSSALPGPPLPARGHALLRWTAALLCLMV 419
 Db 336 -----TMMHICHSPSA-----QYFAIAMLMP 358
 QY 420 VSAFSGFGPVTWLVSEIYVPEIRGAFAP-CN-SFNAANLFISSLPLDGLIGTIGLSW 477
 Db 359 IVCFANSAGPLIWLVCSEIQP--LGRDRFGITCSTATNIANNIVGATFLTMLNLGNAN 416
 QY 478 TLLVGLTAVLGLGYLYFVPEPKGSLAEIDQOFKRR 516
 Db 417 TFWVYAGLVLFILLTLVLPVETKHSVLSHEIERNLMKGR 455
 RESULT 15
 B26430
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 24-Sep-1999
 C:Accession: B26430; A28075; I40996; B65067
 R:Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
 Nature 325, 641-643, 1987
 A:Title: Mammalian and bacterial sugar transport proteins are homologous.
 A:Reference number: A93389; MUID:87115869
 A:Accession: B26430
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-472 <MAI>
 R:Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
 J. Biol. Chem. 263, 8003-8010, 1988
 A:Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabinose isomerase.
 A:Reference number: A28075; MUID:88228015
 A:Accession: A28075
 A:Molecule type: DNA
 A:Residues: 1-472 <MA2>
 A:Cross-references: GB:J03732; NID:gl45320; PIDN:AAA23469.1; PID:gl45321
 R:Stoner, C.; Schleif, R.
 J. Mol. Biol. 171, 369-381, 1993
 A:Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, transcription, and overexpression.
 A:Reference number: I40996; MUID:84114868
 A:Accession: I40996
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-25, 'Y', 349, 'R' <RES>
 A:Cross-references: EMBL:X00272; NID:g40940; PIDN:CAA25075.1; PID:g40941
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: B65067
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-472 <BLAT>
 A:Cross-references: GB:AE000368; GB:U00096; NID:g2367165; PIDN:AAC75880.1; PID:g1789207
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: araE
 A:Map position: 61 min
 C:Superfamily: glucose transport protein
 C:Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane protein

Query Match 17.9%; Score 495; DB 2; Length 472;
 Best Local Similarity 25.2%; Pred. No. 1.2e-25;
 Matches 137; Conservative 93; Mismatches 182; Indels 132; Gaps 8;

QY 17 LGGTFYGLAVISGALLPLQDLGSLCEQFLVGLSLLGALLASLVGGFLIDCYGRKO 76
 Db 31 VAGLLFGLDIGVAGALPFDHFTLSRLQEWVYSSMMLGAAIGALFNGLWLSFLGRKY 90

QY 77 AILGSNLVLLAGSLTLGIAGSLAWLVLGRVAVGFATLSSMACCIYVSELVGPQRGVLV 136
 Db 91 SLNAGAILFVLGSGSFAFATSEVEMLIAARVVLGIVAGIASYTAPLYLSEMASENVRCKMI 150

QY 137 SLYEAGITWGILLSVALNYALAGTTPGWRHMGWATAPAVLOSLSLLFLP----- 186
 Db 151 SMYQLMTVLGIVLAFSLDTAFSYS-GNWRAMLGVLAIPAVLLIILVWFLPSPRWLAEGK 209
 QY 187 -----AGTDETA-----THKDLIPLAGGAPKLGPRPRYSFUDLFRARDNM 228
 Db 210 RHIEAEVLRMLRDTSEKAREELNEISLKKOGGWA-----LFIKINRV 255
 QY 229 RGRITVGLGLVLPFOOLGTOPNLCYASTIFSSVGHGSSAVLASVGLGAVKVAATUTAM 288
 Db 256 RRAVELGMLLQAMQOFTGMNIIMYAPRIKMGAGTTTECOQMIATLVVGLTFEATFIIV 315
 QY 289 GLVDRAGRRALLLAGCALMALS--VSGIGLVSEAVPMDSGPSCSLAVPNATGOTGLPCDSG 346
 Db 316 FTVDKAGRKPAKLTGFSVMALGTLVLGYCLMQF-----DNGTASSG----- 356
 QY 347 LLQDSSLPPIPRTNEDQREPISTAKTKPHRPSGDPSPAPRLALSSALPGPPLPARGHA 406
 Db 357 ----- 356
 QY 407 LLRWTALLCLMVFSASFSGFGPVTWLVSEIYVPEIRGAFAPCNFSFNAANLFISSLF 466
 Db 357 -LSWLSVGMTMMCIAGVMSAAPVWILCSEIQPLKCRDFGTCSTTTNWSNNIIGATF 415
 QY 467 LDLIGTIGLSWTPLLYGLTAVLGLGYLYFVPEPKGSLAEIDQOFKRRFTLSFGHRQN 526
 Db 416 LTLDSIGAAGTFWLYTALNIAFVGITFWLIPETKNVTLIEHIERK-----LMAGEKLR 468
 QY 527 STGI 530
 Db 469 NIGV 472

Search completed: March 15, 2002, 07:07:47
 Job time: 1463 sec